

(19) 대한민국특허청 (KR)

(12) 공개특허공보 (A)

(51) Int. Cl. ⁴

(11) 공개번호

특2000-0023133

H04N 7/247

(43) 공개일자

2000년 04월 25일

(21) 출원번호 10-1999-0039301

(22) 출원일자 1999년 09월 14일

(30) 우선권주장 98/11495 1998년 09월 15일 프랑스 (FR)

(71) 출원인 톰슨 멀티미디어 브릭크만 게오르그

(72) 발명자 프랑스, 예프-92648 블롱 세덱스, 파이 알폰스 르 갈로 46
길로텔필리프

(74) 대리인 프랑스, 예프-35770, 베른쉬르세이치, 뤼드사튀브리앙 60
문경진, 조현석

실사청구 : 없음

(54) 화상 압축 방법과 그 방법을 실행하기 위한 장치

요약

본 발명은 화상 압축, 특히 MPEG 2 타입의 화상 압축에 관한 것으로, 상기 화상은 그룹(GOP)에 따라 코딩되는데, 여기서 각 그룹은 N 화상과, 내부 모드에 코딩된 I 화상과 내부 화상 I 또는 선행 P 화상의 합수로써 예측되는 P 화상을 포함하는데, 각 P 화상은 선행되고, n 개의 양방향성 예측 B 화상이 뒤에 이어지는데, n은 제로가 될 수 있다. 여기서, $M = n + 1$ 은 그룹의 구조를 나타낸다.

그룹에 따라 코딩된 소스 화상을 특징으로 하는 하나 이상의 매개 변수는 테스트 코딩(70)을 사용하여 결정되고, N 과 M은 이러한 매개 변수에 따라 지정된다. 테스트 코딩의 과정에서, 정의된 값들은 N, M값이나, 양자화 인터벌 Q를 지정한다.

도면

도 1

도 2

도 3

도 1은 4.2.0 표준 마크로 블록도.

이를 위해서, MPEG 2 표준은 주기적으로 화상이 공간 모드 즉, 모드에 따라 공간 여유도만을 이용하는 공간 모드에서 코딩 되어 하는 것을 제공한다. 공간 모드로 코딩된 화상은 INTRA 화상 또는 I 화상이라고 한다.

시간 여유도를 이용하여 코딩된 화상은 두 가지 형태가 있는데, 하나는 선행 예측을 기초로 한 일시적인 이전 화상을 기준으로 구성된 화상과, 다른 하나의 형태는, 선행 예측과 후속 예측을 기초로 하여 두개의 일시적인 이전 화상과, 후속 화상을 기준으로 구성된 화상이다.

선행 예측을 기초로 구성된 코딩 화상은 예측 화상이라 불리는 P 화상과, 선행 및 후속 예측을 기초로 하여 구성된 코딩 화상은 양방향 화상이라 불리는 B 화상이다.

I 화상은 제한 없이 그 자체 외의 화상으로 디코딩된다. P 화상은 P 또는 P를 선행하는 I 화상으로 디코딩된다. B 화상은 I 또는 I 화상을 선행하는 B 화상으로 디코딩되고, I 또는 P 화상은 B 화상 다음에 오게 된다.

I 화상의 주기성은 대개 GOP("화상 그룹")으로 표기되는 화상 그룹으로 정의한다.

단일 GOP 내에서, I 화상에 포함된 데이터의 양은 일반적으로 P 화상에 포함된 데이터의 양보다 더 크고, P 화상에 포함된 데이터의 양은 일반적으로 P 화상에 포함된 데이터의 양보다 더 크다.

50 Hz에서 GOP는 I 화상으로 나타나고, 그 뒤를 이어 B 와 P 화상이 따르며, 대부분의 경우에 다음의 시퀀스가 나타난다.

I, B, B, P, B, B, P, B, B

그렇지만, 표준은 $N = 12$ 인 화상이 GOP에 제공되는 것을 요구하는 것은 아닌데, 일반적인 경우에는 두개의 P 화상 사이의 거리 M 은 항상 3이 되지 않는다. 더욱 상세히 말하자면, 거리 M 은 선행 또는 P 화상이 뒤따르는 B 화상의 수 n 이고, 이것은 한 유닛씩 증가되는데, 다시 말해, $M = N + 1$ 이다.

N 은 GOP의 크기 또는 길이를 나타내고, M 은 GOP의 구조를 나타낸다.

본 발명은 보상 레벨을 증가 및/또는 코딩의 질을 증가하기 위한 M 및 N 파라미터상에서 실행 될 수 있다는 것에서 시작된다.

활형이 예측코딩하는 기술적 문제

본 발명에 따른 코딩 방법은 하나 이상의 매개 변수는 그룹에 따라 코딩되는 소스 화상을 나타내기 위해 정해지고, 그룹의 길이와 구조는 상기 매개변수에 따라 이루어지는 것을 특징으로 한다.

한 실시예로, 소스 화상을 나타내는 매개 변수는 정의된 값을 N , M 및 양자화 인터벌 Q 에 할당되는 과정에서 테스트 코딩을 돕기 위해 정해진다.

테스트 코딩은 예를 들자면 개방 루프에서 수행된다.

특히, 간단한 한 실시예에서, 매개변수(P_{cost}) 테스트 코딩시에 획득된 P 화상의 특징을 나타내는 매개변수(P_{cost})와, 테스트 코딩시에 획득한 B 화상을 나타내는 매개변수(B_{cost})가 각각 결정되고, 이러한 매개변수는 바람직하게 P와 B 화상의 코딩의 평균 값이 되는 P와 B 화상을 나타낸다. 화상을 코딩할 때의 코스트는 코딩에 필요한 비트의 수(헤더를 포함하고 있는)가 된다.

이 경우에 있어서, N 은 P 화상을 나타내는 매개변수에 따라 형성될 수 있고, M 은 B 화상을 나타내는 매개 변수에 따라 형성될 수 있다.

본 발명이 말하고자 하는 논지에 따라, 여러 형태의 화상 시퀀스에서, 각 시퀀스 타입에 대해, 최상의 N 은 P 화상에 대한 최소 코딩 코스트를 제공하도록 존재하고, 최상의 M 은 B 화상에 대해 최소의 코딩 코스트를 제공하기 위해 존재한다는 것이 주지되며, 여기서의 코스트는 테스트 코딩 동안에 얻어진다. 이러한 시퀀스는 가변 진폭의 이동, 다른 물체, 다른 반사율, 및 다른 콘텐츠에 의해 다른 것로부터 구별된다.

더욱이, 실제로 선형인 관계가 최상의 N 과 P 화상의 산출량(throughput) 사이에 있게 된다는 것이 주지된다. 이와 같이, 실제로 선형인 관계가 최상의 N 과 P 화상의 산출량(throughput) 사이에 있게된다. 여기서, P 와 B 화상의 산출량을 안다면, 최상의 결과를 제공하는 N 과 M 을 계산하기가 쉬어진다.

한 예로, 50Hz, MPEG2 표준에 해당하는 테스트 코딩은 $N = 12$, $M = 3$ 과 $Q = 15$ 로 실행되고, N 과 P 화상의 산출량 사이의 관계는 대략 다음과 같다.

(1) $N = INT$

$$\left[\frac{389000 - P_{cost}}{10000} \right] + 1, 12 \leq N \leq 30$$

M 과 산출량 사이의 관계 또는 코스트, B 화상의 B_{cost} 는 다음과 같다.

(2) $N = INT$

$$\left[\frac{179000 - B_{cost}}{20000} \right] + 1, 1 \leq M \leq 7$$

M 값의 제한을 5까지도 할 수 있다.

상기 식에서, INT 는 정수 부분을 나타낸다.

12와 30 사이의 N 과, 1과 7사이의 M 으로 한정하는 것은 코드를 간단하게 시행할 수 있게 하고, 프로그램-변환 시간을 줄이는데 가능하게 된다. 동일한 목적으로, 다른 한정이나 값 억제도 가능한데, 특히, M 은 GOP에서 상수 M 및/또는 N 의 서브필들이 된다.

한 실시예에서, 각각 취해진 N 과 M 값과 이들의 합은 억제된 값과 비교되지는 않는다면, 계산된 값과 유사하고 비교된 값을 만족시키는 M 과 N 값이 선택된다. 이런 경우에, M 의 값이 선호된다. 다시 말하면, 여러 M 과 N 쌍 사이에서 선택해야 한다면, 쌍은 계산 결과와 유사하게 얻어진 M 값이 선택될 것이다.

위의 식 (2)는 B_{cost} 가 179800을 넘지 않을 때 적용된다. 반대의 경우에, 말하자면, $B_{cost} > 179000$ 이라면, M 이 다음과 같은 방식 즉,

(3) $M=5, INT$

$$\left\lfloor \frac{P_{cost}}{B_{cost}} \right\rfloor \left\lfloor \frac{M}{5} \right\rfloor$$

선택되어야 함을 보여준다.

B 화상의 코스트가 P 화상의 코스트보다 더 높다면, GOP에 B 화상이 포함되지 않는 것이 바람직하며, 다시 말해, $M = 1$ 이 된다. 이것은 B 화상보다 더 나은 예측 품질을 보여 주고 적은 코스트라고 가정되는 P 화상으로 인해, B 화상의 존재함은 이러한 경우에 있어서 단점이 된다.

비트에서의 각 P 화상과 각 B 화상의 코스트는 예컨대, 이러한 화상이 나타날 때 결정된다. 한 실시예에서, M과 N의 값은 테스트 코딩의 P와 B 화상에 대한 평균값을 위해 선택되고, 코딩은 N 소스 화상의 테스트 코딩 후에만 수행되고, N은 P 화상의 코딩 코스트에 의해 결정된다. 이러한 경우에, 매개변수 M은 GOP에서 상수를 유지하게 된다.

다른 실시예에서, 장면 내용의 변화에 더 빠르게 적응하고, 소스의 화상이 도착하고 진정한 의미의 코딩 사이의 지연 시간을 감소시키는 것을 허용하며(따라서, 저용량 버퍼 메모리가 사용된다) 진정한 의미의 코딩은 테스트 코딩이 데이터가 상기 개시를 허용하는 것을 제공하자마자 개시된다. 테스트 코딩의 제 1 B 화상이 코딩 개시를 허용하는 M을 제공하고, N은 테스트 코딩의 제 1 P 화상에 의해 공급된다. 또한 코딩이 제 1 P 화상이 테스트 코딩을 하고 난 뒤에서 만 시작되게 하는데, 이 경우에, 코딩은 N과 M의 값이 알려질 때 시작된다.

"은 더 플라이" 코딩하는 형태에 있어서, 구조를 말하는 M은 GOP 내에서 변할 수 있는데, 이것은 장면의 내용에 더 빠르게 적응 될 수 있다.

이전에 수행된 코딩에서, 현재의 GOP에 이미 코딩된 화상의 수가 측정된 수 N(위의 예에서 Pcost에 의해 측정된)과 적어도 동일하거나, 장면의 변화가 있을 때 인터럽트 된다.

다음에 나오게 되는 그룹들 사이의 매개 변수에서의 현저한 변화를 피하기 위해서, 계산된 값으로부터 분리되 나오는 것이 가치 있다는 것을 증명할 수 있다. 예컨대, 계산치가 GOP 길이의 큰 부분, 예컨대 80 %에 대해 $M = 1$ 이 될 필요하다는 것을 보여주고, 여기서, 계산치가 GOP의 나머지 부분들에 대해 다른 값이 필요하다는 것을 보이더라도, GOP의 나머지 부분에 대해 계산치는 M은 1보다 더 커야 하고, 모든 값에 대해 1이 되도록 M이 채택된다.

이와 같이, 선택 GOP에 대해, $M = 1$ 이고 현재의 GOP에 대해 계산치가 값 $M = 1$ 은 현재의 GOP의 현저한 부분, 예컨대 60 %에서 필요하다는 것을 보인다면, 계산치가 식 (2)의 결과에 의해 다른 값에 적용되더라도, 값 1은 M에 채택될 것이다.

장면의 변화가 발생할 때, 즉 비디오 화상의 시퀀스에 불연속이 발생할 때, 불연속의 양 부분에 GOP 화상 그룹을 채택하는 것이 필요한데, 이것으로 I 화상으로 개시되는 새로운 그룹이 새로운 장면에 해당하게 된다.

한 실시예에서, 장면의 변화가 그룹에서 일어났다면, 새로운 장면은 새로운 그룹의 I 화상을 구성하고, 영향받은 그룹은 새로운 장면 전에 정지되도록 짧아질 것이다. 만약 장면의 변화가 영향받은 그룹에서 일어났다면, 시작부와의 거리는 허용 가능한 최소 값과 적어도 같게 될 것이다. 영향받은 그룹의 개시부는 영향 받은 그룹의 장면의 변화를 선행하는 화상의 수와 선행하는 그룹의 화상의 수의 합이 최대 허용 가능 N을 넘지 않을 때, 선행하는 그룹을 연장시키는데 사용된다. 이러한 선행 그룹은 변경되는데(짧아지거나, 길어진다), 이것은 상기 GOP에 대해 이전에 계산된 수를 변경하는데 필요하다.

영향받은 그룹의 길이가 N의 최소 허용치보다 더 작은 경우에 바람직하게 사용되는 하나의 변형에서, 그룹에서 장면의 변화가 발생하면, 새로운 장면은 새로운 그룹의 I 화상을 구성하는데, 상기 새로운 그룹은 영향받기 전의 그룹의 길이와 선행되는 그룹의 길이의 평균과 동일한 길이를 갖는다. 이러한 변형으로, GOP를 측정하기 이전에 M을 변경하는 것이 필요하다.

두 개의 변형이 가능한데, 예컨대, 영향받은 그룹의 길이가 최소 허용 가능 N보다 작을 때, 각 변형에 대해, 변형 전에 M, N 작거나 획득된 작 (M, N)의 길이를 계산하고, 가장 거리가 짧은 쌍을 선택함으로써 두개의 변형 사이에서 하나를 선택한다.

파라미터 N , M 을 결정하기 위해, 산출량 측정을 제외한 파라미터측정을 해야한다. 예컨대, N 을 결정하기 위해서, 인트라 화상 I 에 대한 에너지가 사용된다. 이등 진폭이나 이등 보상 에러를 측정하는 것도 가능한데, 이것은 M 과 N 을 측정하기 위한 DFD(치환 프레임 차: Displaced Frame Difference)로 알려져 있다.

본 발명의 다른 특색과 장점은 여러 실시예를 통해 나타낼 것이고, 첨부된 도면을 참조해 언급된다.

발명의 구성 및 작용

먼저는 MPEG2 코딩에 사용되는 특정 원리를 보여주기 위해 도 1에서 도 3을 참조하여 기술한다.

MPEG2 표준에서, 개시 지점은 순환 모드에서 각 720 포인트에 576 라인을 포함하는 화상이 된다. 인터레이스(interlaced) 모드에서는, 상기 화상은 각 720 포인트에 288라인을 포함하는 두 개의 프레임으로 구성된다.

각 화상은 정사각형 $16 * 16$ 휘도 포인트의 제공으로 형성되는 매크로블록으로 분류된다. 따라서 각 매크로블록은 4개의 정사각형 $8 * 8$ 휘도 포인트의 블록으로 형성된다. 이러한 각 4개의 휘도 블록과, 두개의 색상 블록에서, 각각은 $8 * 8$ 포인트를 보여주는데, 블록 중 하나는 신호 C_r 또는 레드의 색상 차를 나타내고, 다른 블록은 C_b 또는 블루의 색상 차를 나타낸다. 4.2.2 포맷에서, 각 휘도 매크로블록은 네 개의 $8 * 8$ 색상 블록, 두 블록은 블루에 대해, 두 블록은 레드와 각각 연관된다. $8 * 8$ 블록을 포함하는 각각의 휘도와 색 성분에 대한 4.4.4포맷도 보여지고 있다.

네 개의 $8 * 8$ 휘도 블록이 도식되는데, 10으로 언급되는 전체와, 각각의 블루 및 레드에 대한 $8 * 8$ 색 블록 (12, 14)이 도식되고, 전체적으로는 4.2.0 표준의 매크로 블록을 도식한 것이다.

각 블록은 휘도 블록(예를 들자면)을 공간 주파수를 나타내는 계수 블록으로 변환 가능한 플렉스 코사인 변환인 DCT로 표기된 변환을 사용하여 코딩된다. 도 2에서 도식되는 바와 같이, 소스 블록(16)은 $8 * 8$ 계수 블록 18로 변환된다. 블록(18)의 상부 좌측 코너(20)는 제로 공간 주파수에 해당하고(블록의 평균 값) 원점(20)으로부터, 수평 주파수는 화상 표(22)에서 나타난 바와 같이, 우측으로 증가하고, 수직 공간 주파수는 화상표(24)에 나타난 바와 같이, 상부로부터 시작하여 하부로 증가한다.

각 매크로 블록에 대해, 코딩의 형태는 "인트라(intra)" 또는 "인터(inter)" 둘 중의 하나가 선택 되어 한다. 인트라 코딩은 DCT 변환을 화상의 소스 블록으로 공급하게 구성되어 있고, 반면, 인터 코딩은 DCT 변환이 소스 블록과 예측 블록 사이의 차를 나타내는 블록으로 공급되는 것을 포함한다.

매크로 블록이 속하는 화상의 형태에 부분적으로 따라 선택된다. 이러한 화상은 세가지 형태가 있는데, 첫번째 형태로 I 또는 인트라로 알려진 형태인데, 코딩은 모든 매크로 블록에 대한 인트라가 된다.

두 번째 형태로, P 또는 예측 형태다, 화상의 이러한 형태에서, 각 매크로블록의 코딩은 인트라 또는 인터 중 하나가 된다. P 타입 화상에 대한 인터 코딩의 경우에, DCT 변환이 상기 P 화상의 현재 매크로 블록과 선행 I 또는 P 화상으로부터 야기되는 예측 매크로블록 사이의 차에 공급된다.

화상의 세번째 형태는 B 또는 양방향형이라고 불린다. 이러한 화상의 형태의 매크로 블록 각각은 인트라 또는 인터 모드로 코딩된다. 인터 코딩은 또한 상기 변환을 B 화상의 현재 매크로블록과 추측 매크로블록 간의 차에 제공한다. 상기 예측 매크로블록은 선행 화상이나 후속 화상 또는 둘 다의 화상(양방향형 예측)으로부터 발생할 수 있으며, 선행 또는 후속이라 불리는 예측 화상은 I 또는 P 형태가 될 수 있다.

화상의 세트는 도 3에서 도식되는데, 다음의 시퀀스 $B, B, P, B, B, P, B, B, P, B, B$ 에 따라 12개의 화상, 즉, 11개의 B 와 P 화상이 오는 I 화상을 포함한다.

gop는 길이 즉, 한 예로 12와 30 사이에 있을 수 있는 다수의 N 길이와, 두개의 P 화상 사이의 거리 즉, 두개의 연속되는 P 화상 사이의 한 유닛씩 증가되는 다수의 B 화상의 거리를 나타내는 구조 파라미터 M 을 특징으로 한다. 한 예에서, 상기 파라미터 M 은 3과 같다. 또한 한 예에 의해, 상기 M 은 1(B 화상이 아닌)과 7사이에 있게 된다. 더욱이 M 은 코더를 간단하게 하기 위해 서브-멀티플 N 이 되는 것이 주지된다.

계속해서, 화상은 코딩되고, 코더에서 N 과 M 은 계속 유지된다.

본 발명은 코딩된 화상의 시퀀스에 따라 달라지도록 존재하고 있는 M 과 N 의 최상값을 관찰하는 것에서 기인한다. 이것은 화상 시퀀스가 정의한 것보다 더 크거나 작게 있는지 다소 현저한 이동을 보이고 있는지에 따라, M 과 N 의 최상 값은 현저히 달라질 수 있기 때문이다. 최상 값에 의해 동일한 품질에 있어서, 최소의 비트를 요구한다는 것이 주지된다.

우선 적으로, 본 발명의 요지에 따라 실행되는 실험 연구는 화상의 정의된 시퀀스에 대한 Gop 의 최상 크기 N 은 P 화상(헤더를 포함하고 있는)을 코딩하는데 필요한 수의 비트의 최소 값 $Pcost$ 에 해당한다는 것이 보여진다. 상기 값은 도 4의 횡 좌표에 점으로 표시된 N 과 I 로 표시된 시퀀스에 대한 $Pcost$, $Pcost$ 값은 시퀀스 I 에 대한 평균값의 P 화상을 코딩하는데 사용되는 비트의 수이다. 따라서, $Pcost(I)$ 는 N 의 값이 최상($Nopt$)인 최소(34)를 나타내는 커브(32)에 의해 나타내어 지는 것을 볼 수 있다.

이와 유사한 방법으로, M 의 최상 값은 정의된 시퀀스에 대한 B 화상을 코딩하는 평균값이 사용되는 비트의 수의 최소 값 $Bcost(I)$ 에 해당하며, I 로 기록된다. 도 5에서 M 은 횡좌표에 나타나고, $Bcost(I)$ 는 세로좌표에 나타난다. 커브(36)는 최상 값 $M(Nopt)$ 에 해당하는 최소값(38)을 나타내는 다이어그램에 나타난다. "홀스(horse)", "플라워 가든(flower garden)" 및 "모브컬(Mobcal)"로 표시되는 MPBG 코딩에 사용되는 테스트 시퀀스가 측정된다. "홀스" 시퀀스는 해상도가 좋으면서 이동성이 빠른 것에 해당하고, "플라워 가든" 시퀀스는 또한 평균 이동성에 좋은 해상도에 해당하고, 반면, "모브컬"시퀀스는 높은 해상도와 느린 이동성에 해당한다. 다른 시퀀스들도 테스트되는데, 카약(kayak)시퀀스는 빠른 이동성에 낮은 해상도, 바스켓(basket)시퀀스는 평균적으로, 한 동작과, 좋은 해상도 화상을 갖는 시퀀스들이 테스트된다.

그룹이 M , N 과 양자화 간격(Q)의 정해진 값으로 테스트 코딩된다면, 이러한 값은 시퀀스의 최적 값에 반드시 해당되지는 않고, P 화상 $Pcost$ 를 코딩하는 평균 코스트와 B 화상 $Bcost$ 를 코딩하는 평균 코스트는 각각 N , M 을 나타낸다. 더욱이, 도 6에서, 각 시퀀스 I 에 대한 $Nopt$ 과 소정의 M , N 및 Q 에서의 $Pcost$ 를 코딩하는 사이의 간단한 관계가 형성된다. 이러한 관계는 선형이거나 실제적으로도 선형인데, 직선(40)으로 도시되고 있으며, 다른 포인트(42, 44) 등은 다른 시퀀스를 나타낸다.

도 7은 $Nopt$ 가 횡좌에 있고, 코딩 코스트($Pcost$) (M , N 및 Q 로 정의된)가 세로축에 있는 다이어그램인데, 각 포인트(52, 54, 56) 등은 소정의 시퀀스에 해당한다. 이러한 포인트는 직선(60)에 있게 되는 것이 보여진다. 따라서, $Nopt$ 와 테스트 코딩의 코스트사이에 선형 관계가 있게 된다.

테스트 코딩 과정에서 사용되는 M , N , 및 Q 의 값이 다음과 같을 때,

$$M = 12,$$

$$N = 3 \text{ 및}$$

$$Q = 15$$

M 과 N 값은 다음의 관계를 만족시킨다.

$$(1) N = INT$$

$$\langle \text{MARGIN} \rangle \langle 1/R \rangle \left[\frac{389000 - P_{\text{cost}}}{10000} \right] 11.12345678 \langle 1/P \rangle \langle 1/P \rangle$$

(2) $N = \text{INT}$

$$\langle \text{MARGIN} \rangle \langle 1/R \rangle \left[\frac{179000 - B_{\text{cost}}}{2000} \right] 11.2345678 \langle 1/P \rangle \langle 1/P \rangle$$

비록, 위의 식 (2)에서, M 은 1과 7사이에 있어야 하고, 도 7의 도면에서 보여지며, M 은 실제로 5 까지로 제한된다.

본 발명을 실행하기 위한 설계도가 도 8에 도시된다. 이것은 테스트 코딩 또는 "제 1 패스"코딩을 실행하기 위한 제 1 MPBG 코드 (70)를 포함한다. 상기 테스트 코딩은 위에서 언급된 즉, 예로 $M = 12$, $N = 3$ 및 $Q = 15$ 인 고정된 매개변수로 설정된다. 상기 테스트 코드는 여기의 예에서, 개방 루프에서 동작, 즉 제한 없이 동작하게 된다.

코드 (70)는 도 6과 도 6 및 (1)과 (2)의 관계에 따라 도시된 바와 같이, P_{cost} 를 N_{opt} 로 B_{cost} 를 M_{opt} 로 전환하는 변환기 (72)에 제공되는, B_{cost} 와 P_{cost} 값을 공급한다.

이러한 값 (N , M)은 위에서 기술된 바와 같이 화상 그룹에 대해 계산되고, MPBG 2 코드 (74)의 제어 입력 (76)에 공급된다.

코드 (74)의 입력에서의 데이터는 테스트 코드 (70)의 입력 데이터와 동일하다. 버퍼 메모리 (78)가 테스트 코드 (70)와 변환기 (72)에서의 처리 시간에 따라 공급되고, 상기 메모리 (78)는 처리하는 동안 데이터를 유지한다.

변환기 (72)에서, 식 (1), (2)로부터의 N , M 값은 실제에 부여된 강제성 (constraint)과 비교되는데, 특히, M 은 N 의 서브-멀티플이다. 계산된 값이 비교되지 않을 경우, N 과 M 의 값은 계산된 값과 유사한 값을 채택하지만, 대개는 M 값이 선택된다.

변환기 (72)는 보상 조건에 대해서도 고려할 수 있다.

우선, B_{cost} 와 P_{cost} 사이의 비교를 수행하는데, B_{cost} 가 P_{cost} 보다 더 높을 경우, 값 1이 M 에 할당되고, GOP는 B 화상을 포함하지 않는다. 이것은 이러한 가정에 있어서, B 화상은 P 화상보다 더 높은 코딩 코스트를 갖기 때문에, 높은 예측 품질을 보이는 P 화상만을 유지하는 것이 바람직하다.

두번째로, 변환기는 B_{cost} 를 값 179000과 비교하며, B_{cost} 가 179000 초과하면, 관계 (2)는 다음의 관계로 대체된다.

(3) $M=5, \text{INT}$

$$\left[\frac{P_{\text{cost}}}{B_{\text{cost}}} - 1 \right], 1 \leq M \leq 7$$

변환기 (72)는 또한 두가지 특별한 경우에 고려될 수 있는데, 여기서 화상 품질의 단일성을 얻기 위해 관계 (2)로부터 벗어나는 것이 필요하다.

제 1 경우는 다음과 같다. 테스트 코딩은 M 이 적어도 2와 동일한 값을 보여야하지만, 더욱이 테스트 코딩은 M 에 의해 얻어진 중간 값이 그룹의 큰 부분, 예컨대 60 % (이러한 제한은 제 1 경우의 기대값 이하)에 걸쳐 1과 동일해야하고, 선행 그룹은 $M = 1$ 과 같이 된다. 이러한 경우에, 값 1은 M 에 설정된다.

M 에 대해 값 1로 설정되는 이러한 두 특별한 경우는 연속 그룹에 걸친 시퀀스의 동일한 형태에 대해서 이러한 조건들은

좋은 단일 품질을 허용한다는 것을 보여주는 본 발명의 의도에 따라 실행되는 테스트로부터 얻어질 수 있다.

마지막으로, 변환기(72)는 장면의 변화 또는 코더에서 대개 검출되는 "컷(cut)"을 고려한다. 이러한 장면의 변화가 발생 할 때, GOP는 새로운 장면으로 시작되는데, 다시 말해, 새로운 장면이 나오면, 인트라 I 화상이 공급된다.

더욱이, 본 발명의 방법에 따라, 장면의 변화가 검출될 때, 선행 GOP와 현재의 GOP는 다음의 조건을 기초로 해서 구성 된다.

장면의 변화가 12번째 화상 다음의 GOP에 나타난다면, 새로운 GOP는 이때, 장면의 변화를 시작하고, 따라서, 선행 GOP는 제한되거나, 짧아진다.

대조적으로, 장면의 변화가 12번째 화상 전에 나타난다면, 선행 GOP를 제한하지 못하게 되는데, 장면의 변화 이전에 끝 내기 위해 이때, 선행 GOP를 제한하지를 못한다면, 이 경우에, 화상의 수는 요구된 최소 수보다 더 작게 된다. 선행 GOP와 현재의 GOP는 이때 다음의 방법으로 변경되는데, 두 가지 경우로 구별된다.

첫 번째 경우로, 장면의 변화는 선행 GOP의 화상의 수와 현재 GOP의 화상의 수의 합이 장면의 변화 전에 최소한 30과 동일하게 되는 순간에 나타난다. 이 경우에, 선행 GOP는 길어진다.

두 번째 경우에서, 장면 변화 전 화상에 해당하는 현재 GOP 화상의 수와 선행 GOP 화상 수의 합은 30 보다 크다. 그리고 선행 GOP와 현재 GOP는 상기 두개의 GOP에 대한 평균을 계산하여 재배치된다.

예컨대, 만약에 선행 GOP에서 $N = 25$ 이고 $M = 2$ 가 되고, 또한 계산치가 $N = 20$ 이고 $M = 3$ 인 현재 GOP의 8번째 화상 뒤에 장면 변화가 발생한다면, 길이가 짧아진 현재 GOP만큼 길이가 길어진 선행 GOP는 33개의 화상을 포함한다. 상기 값이 최대 허용치(30)를 초과하게 되면, 전체 화상 수가 33인 두 개의 GOP에 대한 "평균"을 은 두개의 GOP에 해당하게 되고, 이것의 전체 수는 33이 되며, 각각의 GOP는 부여된 억제값 비교되어야 한다. 이 경우에, 선행 GOP에 대해, $N = 19$ 과 $M = 2$ 사이에서 선택이 이루어지고, 장면이 변하기 전에 $N = 15$ 및 $M = 3$ 사이에서 선택이 이루어짐을 볼 수 있다. 길이가 19과 15는 선행 그룹(25)의 길이와 영향 받은 현재 그룹의 길이(8)의 평균(16, 5)에 근접한다.

테스트는 장면의 변화, 빛의 플래시 및 상대적으로 긴 지속 기간 및 일반적인 코딩 방법으로 얻어지는 결과물을 갖는 12 개의 서로 다른 시퀀스에 대해 실행되는데, M과 N의 시퀀스의 값에 채택되는 발명에 따른 방법으로 얻어지는 결과를 비교하는 M, N의 고정된 값에 해당한다. 이러한 테스트는 여러 산출량으로 실행된다. 품질의 향상이 주지되는데, 이것은 0.2dB에서 1.14dB의 PSNR 매개변수(피크 신호에 대한 잡음 비율; Peak Signal to Noise Ratio)에 의해 측정된다. PSNR에서의 이러한 증가는 2에서 22 사이에 있는 비트에 대한 저장에 해당한다.

본 발명에 따른 방법은 임의의 형태의 비디오 화상 압축 방법에서 사용될 수 있는데, 여기서 I, P 및 B 화상이 제공된다. 이것은 실시간에서나, 오프라인에서의 레코딩과 전송 둘 다에 적용될 수 있다.

상기 방법은 GOP의 크기가 코딩되기 전에 결정되는 경우에 대해서만 제한되는 것은 아니다. 이것은 또한 파라미터M과 N이 각 화상에 대해 계산 될 때에도 적용될 수 있는데, 적절한 코딩은 온 더 플라이로 실행된다. 이 경우에 있어서, M은 GOP내에서 실행될 수 있는데, 예를 들자면, 현재의 GOP에서 코딩되는 화상의 수가 계산된 N과 적어도 동일할 때, 개시되는 새로운 GOP 내에서 실행될 수 있다. M은 GOP 내의 화상의 복잡도에 관한 함수로써 변할 수 있다.

이러한 경우에서, 버퍼 메모리(78)(감소되는 용량)의 GOP 전체를 저장할 필요는 없는데, MPEG 2 표준에서만 표시된다면, M과 N의 제한 값이 줄어들게 된다. 장면의 변화에 부여된 제한은 그리 심하지는 않다.

본 발명은 이하에 나오는 청구범위를 범위를 벗어나지 않는 범위 내에서 위에서 제시한 실시예에 제한되지 않고 다른 많은 변형된 실시예가 가능하다는 것은 당업자들은 알 수 있다.

(S7) 청구항 해석

청구항 1. 픽처 그룹(GOP) 화상이 코딩될 때의 화상 압축 방법으로서, 상기 GOP는 N 개의 화상을 포함하는데, 상기 N 은 상기 그룹의 길이를 표시하며, 또한 상기 GOP는 필드내 부호화 모드로 부호화된 I 화상과 필드내 부호화 모드로 부호화된 상기 I 화상이나 선행하는 p 화상에 기초하여 예측된 p 화상을 포함하는데, 상기 p 화상 각각은 n 개의 양방향 예측 화상에 선행하는데 즉 한 개의 p 화상 뒤로 n 개의 B 화상이 쫓아 나오며, 상기 n 은 어쩌면 제로 일 수 있고, 상기 M 은 상기 n 에 1을 더한 것으로서 그룹의 구성을 표시하는, 상기 화상 압축 방법에 있어서,

그룹 단위로 부호화될 소스 화상의 특성을 나타내는 한 개 이상의 파라미터가 결정되고, 또한 그룹의 길이(N) 및 구성(M)은 상기 파라미터에 의거하여 만들어지는 것을 특징으로 하는 화상 압축 방법.

청구항 2. 제 1 항에 있어서, 상기 소스 화상의 특성을 나타내는 상기 파라미터가 소정의 값에 N , M 및 양자화 인터벌(Q)에 할당되는 과정에서의 테스트 부호화를 통해서, 결정되는 것을 특징으로 하는, 화상 압축 방법.

청구항 3. 제 2 항에 있어서, 상기 테스트 부호화는 개방 루프로 실행되는 것을 특징으로 하는 화상 압축 방법.

청구항 4. 제 2 항 또는 제 3항에 있어서, 상기 소스 화상의 특성을 규정하기 위해서, 상기 테스트 부호화가 이루어지는 동안에 얻어진 p 화상의 특성을 나타내는 파라미터($Pcost$)와 상기 테스트 부호화가 이루어지는 동안에 얻어진 상기 B 화상의 특성을 나타내는 파라미터($Bcost$)가 결정되는 것을 특징으로 하는, 화상 압축 방법.

청구항 5. 제 4항에 있어서, 상기 수(N)은 하나 이상의 p 화상의 특성을 나타내는 파라미터에 기초하여 결정되고, 상기 수(M)은 하나 이상의 B 화상의 특성을 기초로 하여 결정되는 것을 특징으로 하는, 화상 압축 방법.

청구항 6. 제 4항 또는 제 5항에 있어서, 상기 p 화상과 B 화상의 특성을 나타내는 상기 파라미터($Pcost$, $Bcost$)는 여려단 평균 비용과 같은, 상기 p 화상과 B 화상의 부호화 코스트인 것을 특징으로 하는, 화상 압축 방법.

청구항 7. 제 5항 또는 제 6항에 있어서, 상기 테스트 부호화 중에, 각 B 화상을 부호화하는데 드는 평균 비용($Bcost$)이 각 p 화상을 부호화하는데 드는 상기 평균 비용($Pcost$)보다 더 클 때, M 에 1이 부여되고, 그래서 픽처 그룹에는 B 화상이 포함되지 않는 것을 특징으로 하는, 화상 압축 방법.

청구항 8. 제 6항 또는 제 7항에 있어서, 상기 테스트 부호화 중에, 각 B 화상과 해당 M 의 부호화 코스트는 상기 소스 화상의 도착 단계에서 결정되는 것을 특징으로 하는 화상 압축 방법.

청구항 9. 제 8항에 있어서, 상기 테스트 부호화가 끝나기 전에 결정된 상기 M 은 상기 픽처 그룹에서의 대부분에서 1과 동일하고, 상기 그룹의 M 에 값 1이 부여되는 것을 특징으로 하는, 화상 압축 방법.

청구항 10. 제 8항 또는 9항에 있어서, 상기 테스트 부호화가 끝나기 전에 결정된 상기 M 은 상기 그룹의 하나 이상의 제한된 부분에 대해서 1과 동일하고, 상기 M 이 상기 선행 그룹에 대해 1과 동일 할 때, 상기 그룹의 M 에 값 1이 부여되는 것을 특징으로 하는, 화상 압축 방법.

청구항 11. 이전의 항 중 어느 한 항에 있어서, 그룹에서 장면의 변화가 발생하는 경우에, 상기 새로운 장면은 상기 새로운 그룹의 I 화상을 포함하고, 상기 영향 받은 그룹은 상기 장면의 변화가 상기 영향 받은 그룹에서 일어날 경우에, 상기 새로운 장면 전에 장면이 정지될 수 있게 짧아지고, N 에 허용될 수 있는 최소 수와 적어도 동일한 시작부에서부터의 거리에서, 상기 영향받은 그룹의 시작부는 상기 영향받은 그룹에서의 장면의 변화를 선행하는 화상의 수와 허용 가

능한 최대치 N 을 넘지 않는 그룹의 화상 수의 합이 최대치를 넘지 않을 때 상기 장면을 선행하는 그룹을 짧게 하는데 사용되는 것을 특징으로 하는, 화상 압축 방법.

청구항 12. 제 1항 내지 제 9항 중 어느 한 항에 있어서, 그룹 내에서 장면의 변화가 발생할 경우에, 상기 새로운 장면은 새로운 그룹의 I화상을 포함하고, 상기 영향받은 그룹과 선행하는 그룹은 선행하는 그룹과 변경 후의 그룹의 길이와 평균과 근접하는 길이가 되도록 배열되는 것을 특징으로 하는, 화상 압축 방법.

청구항 13. 제 6항에 있어서, $N = 12$, $M = 3$, $Q = 15$ 인 50Hz, MPBG 형태로 실행되는 테스트 부호화에서, N 과 M 은 각각,

$$(1) N = INT$$

$$\left\lceil \frac{389000 - P_{cost}}{10000} \right\rceil + 1, \quad 12 \leq N \leq 30$$

$$(2) M = INT$$

$$\left\lceil \frac{179000 - B_{cost}}{20000} \right\rceil + 1, \quad 1 \leq M \leq 7$$

에 따르는 B 및 P 의 평균 부호화 코스트의 함수인 것을 특징으로 하는, 화상 압축 방법.

청구항 14. 제 13항에 있어서, $1 \leq M \leq 7$ 인 것을 특징으로 하는, 화상 압축 방법.

청구항 15. 제 13항 내지 제 14항에 있어서, 상기 코스트 부호화(B_{cost})가 179000보다 클 때, 상기 M 은

$$(3) M = 5, INT$$

$$\left\lceil \frac{P_{cost}}{B_{cost}} - 1 \right\rceil, \quad 1 \leq M \leq 7$$

에 따라 결정되는 것을 특징으로 하는, 화상 압축 방법.

청구항 16. 제 1항 내지 제 6항 중 어느 한 항에 있어서, 상기 M 은 그룹 내에서 변하는 것을 특징으로 하는, 화상 압축 방법.

청구항 17. 제 2항 내지 제 6항에 있어서, 상기 압축은 상기 테스트 부호화 후에 실행되는 것을 특징으로 하는, 화상 압축 방법.

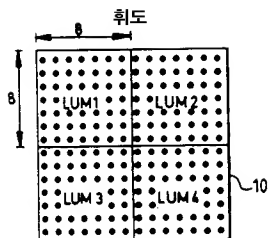
청구항 18. 제 4항 내지 제 6항 중 어느 한 항에 있어서, 상기 압축은 제 1 B 화상 또는 제 1 P 화상을 나타내는 파라미터가 결정된 후에, 시작되는 것을 특징으로 하는, 화상 압축 방법.

청구항 19. 제 18항에 있어서, 부호화된 그룹의 형태는 코딩된 화상의 수가 상기 현재의 P 화상을 기초하여 결정된 N 과 적어도 동일한 것을 특징으로 하는, 화상 압축 방법.

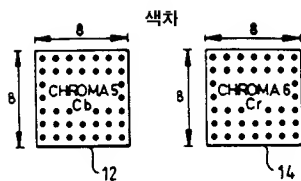
청구항 20. 제 4항 내지 제 6항 중 어느 한 항에 있어서, 상기 테스트 부호화를 수행하고, M 과 N 파라미터를 결정하기 위한 채널을 포함하는 채널과, 상기 부호화를 적절하게 수행하기 위해 제 1 채널로부터 정보를 수신하는 부호화 채널(79, 74)을 포함하는 것을 특징으로 하는, 화상 압축 방법.

도 1

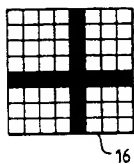
도 2



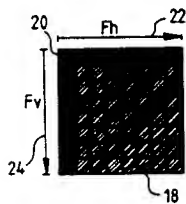
도 3



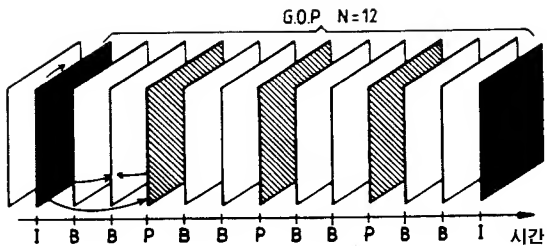
소스 블록



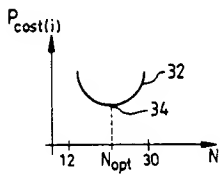
DCT



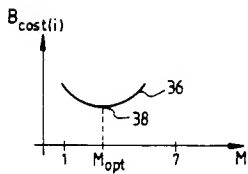
도 4



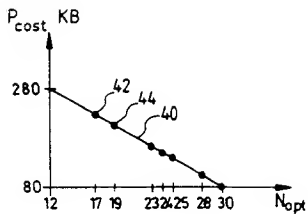
32
34



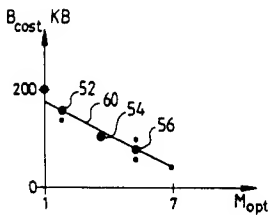
36
38



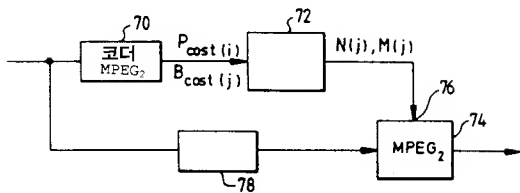
32
34



도 23



도 24



02 - BR 9904108
04 - CN 1248864
07 - EP 987903
10 - FR 2783388
15 - JP 2000102021
20 - KR 2000023133
27 - ZA 9905802

The method involves coding images according to groups each of which comprises N images, with an I image coded in intra mode, P images predicted as a function of the intra image I or of the preceding P image. Each P image is preceded and followed by n bidirectionally predicted B images, n possibly being zero. The number $M = n + 1$ represents the structure of the group.

At least one parameter characterising the source images which are to be coded according to a group is determined using a test coding (70). The numbers N and M are made to depend on this parameter. In the course of the test coding, defined values are conferred on N, M and on the quantisation interval Q.